Neuropsychologia 46 (2008) 2556–2562

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Assessing age-related multisensory enhancement with the time-window-of-integration model

Adele Diederich[a,*], Hans Colonius[b], Annette Schomburg[a]

[a] *Jacobs University Bremen, Germany*
[b] *Department of Psychology, University of Oldenburg, P.O. Box 2503, D-26111 Oldenburg, Germany*

ARTICLE INFO

Article history:
Received 27 November 2007
Received in revised form 18 March 2008
Accepted 31 March 2008
Available online 12 April 2008

Keywords:
Multisensory integration
Time-window-of-integration
Saccadic eye movement

ABSTRACT

Although from multisensory research a great deal is known about how the different senses interact, there is little knowledge as to the impact of aging on these multisensory processes. In this study, we measured saccadic reaction time (SRT) of aged and young individuals to the onset of a visual target stimulus with and without an accessory auditory stimulus occurring (focused attention task). The response time pattern for both groups was similar: mean SRT to bimodal stimuli was generally shorter than to unimodal stimuli, and mean bimodal SRT was shorter when the auditory accessory was presented ipsilaterally rather than contralaterally to the target. The elderly participants were considerably slower than the younger participants under all conditions but showed a greater multisensory enhancement, that is, they seem to benefit more from bimodal stimulus presentation. In an attempt to weigh the contributions of peripheral sensory processes relative to more central cognitive processes possibly responsible for the difference in the younger and older adults, the time-window-of-integration (TWIN) model for crossmodal interaction in saccadic eye movements developed by the authors was fitted to the data from both groups. The model parameters suggest that (i) there is a slowing of the peripheral sensory processing in the elderly, (ii) as a result of this slowing, the probability of integration is smaller in the elderly even with a wider time-window-of-integration, and (iii) multisensory integration, if it occurs, manifests itself in larger neural enhancement in the elderly; however, because of (ii), on average the integration effect is not large enough to compensate for the peripheral slowing in the elderly.

1. Introduction

Numerous studies, both neurophysiological and behavioral, have shown that combining information from different senses can have a profound influence on our perception (for a review see, Calvert, Spence, & Stein, 2004; Stein & Meredith, 1993). The ventriloquist effect (Howard & Templeton, 1966; Woods & Recanzone, 2004) or the McGurk effect (McGurk & MacDonald, 1976) are prominent examples for how visual and auditory information interact in a highly complex manner. *Crossmodal* or *intersensory interaction*, as it is called, can also be observed for simpler tasks such as detection or discrimination, with relatively simple stimuli such as monochromatic lights and pure tones. Todd (1912) was the first to apply a *focused attention* task in which a participant was instructed to respond to a visual stimulus upon detection and to ignore all other stimuli, such as tone or tactile stimulation, that might accompany the light. He found that the reaction time (RT) to bimodal or trimodal stimuli was on average faster than to the visual stimulus when presented alone. This result has been replicated many times, including all kinds of variations of the original setup (e.g., Diederich & Colonius, 2004, for review). The effect has also been termed *intersensory facilitation*(see Welch & Warren, 1986, for definitions and examples of other facilitation effects on sensation thresholds or judgments).

Participants in these experiments have predominately been college-aged adults. Recently, Laurienti, Burdette, Maldjian, and Wallace (2006) compared the performance of older and younger adults in a speeded discrimination task with unimodal (red and or blue filled circle) and bimodal stimuli (circle accompanied by the spoken word *red* or *blue*). Older adults were significantly slower in both unimodal conditions and the crossmodal condition when compared with younger adults. However, the former group showed greater intersensory facilitation than the younger adults, reiterating a finding from more complex tasks that the elderly population seems to benefit more from multisensory input than do

---

* Corresponding author at: School of Humanities and Social Sciences, Jacobs University Bremen, P.O. Box 750 561, D-28725 Bremen, Germany. Tel.: +49 421 200 3431; fax: +49 421 200 3303.

E-mail addresses: a.diederich@jacobs-university.de (A. Diederich), hans.colonius@uni-oldenburg.de (H. Colonius).

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.03.026

younger adults (e.g., Helfer, 1998). It remains controversial, however, whether this reflects age-related differences in multisensory processing, or whether it is associated with a general cognitive slowing observed in the older age groups (Cerella, 1985). In a study from the same lab, Peiffer, Mozolica, Hugenschmidt, and Laurienti (2007), trying to eliminate most high-order cognitive processing, compared the performance of these groups in a simple detection RT task with green light-emitting diodes (LEDs) as visual stimuli and broadband white noise as auditory stimuli. No significant differences in unimodal response times were seen. However, older adults continued to show significantly greater multisensory enhancement than younger adults. These results support the hypothesis that a general cognitive slowing cannot be the sole source for the multisensory superiority of the older adults.

Multisensory RT enhancement is commonly measured by comparing the observed response time distribution with the distribution predicted by a race model (Colonius, 1990; Colonius & Diederich, 2006; Miller, 1982; Molholm et al., 2002; Raab, 1962). This comparison takes into account possible probability summation effects, i.e., the fact that in a bimodal setting stimuli from either modality have a chance of being processed first. From their race model analysis, Laurienti et al. (2006) concluded that older adults have a broader temporal window of integration than younger adults (Laurienti et al., 2006, p. 1161). The processing mechanism responsible for the different levels of multisensory enhancement in both age groups remains unclear, however. The purpose of the current study was to probe whether the time-window-of-integration (TWIN) model recently suggested by us (Colonius & Diederich, 2004; Diederich & Colonius, 2007a,b, 2008) may shed some light on the issue of the stage at which multisensory processing differs in the old and young adults. Specifically, the TWIN model distinguishes a first, peripheral stage where a parallel independent race among the stimuli from different modalities takes place, from a subsequent stage of multisensory integration. When a parametric version of the model is fitted to data from both age groups separately, the estimated parameter values are expected to reveal whether the groups differ with respect to their peripheral sensory processing speed or their ability to integrate information from different modalities, or possibly both.

Next we report the results from a focused attention study where saccadic eye movements to a visual target stimulus were measured with and without an auditory non-target being presented at an ipsi- or contralateral position to the target. A summary of the TWIN model and an analysis of the data fits follow.

2. Methods

2.1. Participants

Fifteen elderly people (aged 65–75, mean 69.6, six male) from the community served as voluntary participants. They had corrected-to-normal vision. Their hearing ability was according to their age cohort and they were tested prior to being admitted to the experiment. Moreover, six undergraduate students (aged 20–22, three female) served as paid voluntary participants as well. All had normal vision and normal hearing. Participants were screened for their ability to follow the experimental instructions (proper fixation, few eye blinks during a trial, saccades directed towards the visual target). They gave their informed consent to their inclusion in the study. The experiment was conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

2.2. Stimuli and apparatus

Auditory stimuli were bursts of (Gaussian) white noise (59 dBA) generated by two speakers (Canton Plus XS). They were placed at 20° to the left and right of the fixation LED at the same height as the participants' ear level and 120 cm in front of the participants. The visual stimuli were red LEDs (3.3 mcd) positioned on top of the loud speakers. The fixation point (red LED, 5.95 mcd) was placed between the target stimuli at the same height.

2.3. Experimental procedure

Participants were seated in a completely darkened and sound attenuated room with the head positioned on a chin rest and the elbows and lower arms resting comfortably on a table. Every experimental session began with 10 min of dark adaptation during which the measurement system was adjusted and calibrated.

Each trial began with the appearance of the fixation point. After a variable fixation time (700–1200 ms), the fixation LED disappeared and, simultaneously, the visual target stimulus was turned on (i.e., there was no gap). Participants were instructed to gaze at the visual target as quickly and as accurately as possible ignoring any auditory non-targets (focused attention paradigm). The visual target appeared alone or in combination with an auditory non-target in ipsi- or contralateral position. 160 trials were presented in one block of trials, lasting for about 10–15 min.

The onset of the noise burst was shifted by a stimulus onset asynchrony (SOA) of -100, -50, 0, or 50 ms. Negative values mean that the non-target was presented before the target. The visual stimuli were presented for 500 ms; the auditory non-targets were turned off simultaneously with the visual stimulus. Thus their duration varied between 600 and 450 ms, depending on SOA. Stimulus presentation was followed by a break of 2000 ms in complete darkness, before the next trial began, indicated by the onset of the fixation LED.

Including one block of training the younger participants performed three blocks of trials each within one experimental session lasting for about 1 h. Each participant was engaged for about 2.5 h. The elderly participants performed two blocks of trials within one experimental session of about an hour. The calibration procedure took considerably longer than for the younger participants. Moreover, they needed longer rests between two experimental blocks. One block of training and the hearing screening included, each one was engaged for 4 h. One block consisted of 160 trials, 32 unimodal (20%) and 128 bimodal stimuli. Each participant completed a total of 960 experimental trials (96 trials per bimodal stimuli and 192 trials for unimodal stimuli).

2.4. Data collection and saccade screening

Saccadic eye movements were recorded by an infrared video camera system (EyeLink II, SR Research) with a temporal resolution of 500 Hz and a horizontal and vertical spatial resolution of 0.01°. Two interlinked PCs controlled the EyeLink program and a third PC controlled stimulus presentation. Criteria for saccade detection on a trial-by-trial basis were velocity ($<35°$/s) and acceleration ($<9500°/\,s^2$). The recorded eye movements from each trial were checked for proper fixation at the beginning of the trial, eye blinks and correct detection of start and end point of the saccade.

3. Results

Saccades were screened for anticipation errors (SRT $<$ 80 ms), misses (SRT $>$ 500 ms), and accuracy: trials with saccade amplitude deviating more than 1STD from the mean amplitude were excluded from the analysis. This removed about 11% of the saccades for the elderly and 16% for the younger group. Table 1 lists the percentages of different error types for both age groups. The error rates are quite low throughout. In particular, combining the saccades made before any signal appeared and the anticipatory errors (SRT $<$ 80 ms) the rate is less than 1%. The rate of directional errors, i.e., gazing at the direction opposite to the signal, is larger for the elderly (1.66%), possibly due to differences in attention level between both groups. There was no evidence for multiple saccades in the remaining data set.

Elderly participants were considerably slower than younger participants under all conditions. Their mean SRTs ranged from 203 ms to 293 ms compared to 137 ms to 173 ms for the younger participants. Moreover, the standard deviation of SRT across elderly

Table 1
Percentage of errors by type for both age groups

Type of error	Age group 65–75	Age group 20–22
Saccades before any signal	2.23	0.38
Amplitude not within 1STD	11.04	15.39
SRT $<$ 80	0.49	0.91
SRT $>$ 500	0.43	0.01
Directional	1.66	0.05
Total	15.82	16.74

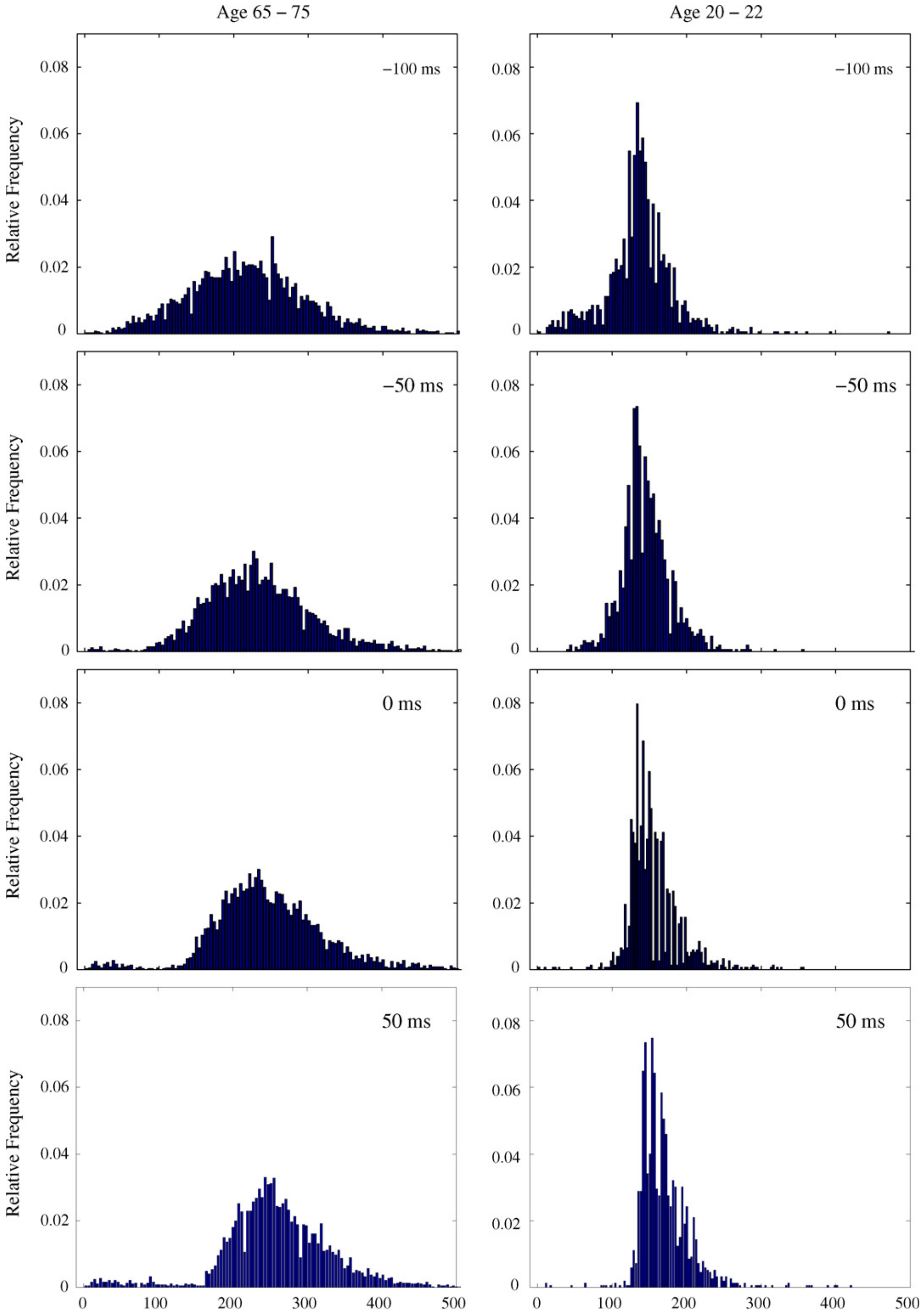

Fig. 1. Histograms of saccadic reaction times for the elderly group (left column) and the younger group (right column) for all SOA values.

participants was about twice as large as for the younger group (64 ms vs. 33 ms). This is also reflected in the histograms in Fig. 1 (left column: elderly group; right column, younger group) where the data were combined across both spatial configurations. The histograms become more and more skewed with increasing SOA values.

The *response time pattern* for both groups is very similar, however: mean SRT to bimodal stimuli was generally shorter than to unimodal stimuli. Furthermore, mean bimodal SRT was shorter when stimuli were presented ipsilaterally than contralaterally (for details see Table 2 containing the means, standard deviations, and standard errors of both groups for the different SOAs and ipsi-and contralateral stimulus configurations). For both groups, the shortening of mean SRT under both spatial conditions is considerable, but even more so for the elderly (up to 90 ms). That is, this group seems to benefit more from bimodal stimulus presentation than

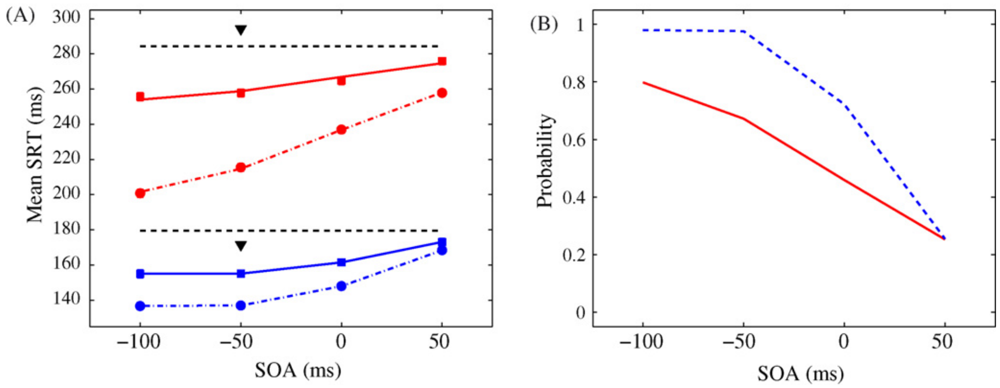

Fig. 2. (A) The fit of TWIN to mean SRT. The upper three curves represent fits and data of age group 65–75, the lower three curves those for age group 20–22. The triangles indicate mean SRT to the unimodal stimulus, i.e., to LED, the circles indicate the mean SRT to bimodal stimuli, i.e., light and sound, presented ipsilaterally, and the squares indicate the mean SRT to bimodal stimuli, when presented contralaterally. Standard errors are smaller than the symbols. The curves represent the predictions of the model. (B) Predicted, but not observable, probabilities of interaction. The solid line represents the predictions for age group 65–75, and the dashed line those for age group 20–22.

Table 2
Mean SRT, standard deviation (STD) and standard error (SE), and multisensory response enhancement MRE $\equiv [(\mathrm{SRT_V} - \mathrm{SRT_{VA}})/\mathrm{SRT_V}] \times 100$ for both age groups

Laterality	SOA	Age 65–75				Age 20–22			
		SRT	STD	SE	MRE	SRT	STD	SE	MRE
	−100	202	66.9	1.68	31	137	33.7	1.41	20
	−50	216	64.6	1.61	27	139	30.1	1.21	19
Ipsi	0	238	63.7	1.55	19	149	29.4	1.19	13
	50	258	59.1	1.47	12	168	31.0	1.26	2
	−100	257	63.6	1.71	13	155	36.1	1.54	10
	−50	258	63.4	1.62	13	156	29.6	1.14	9
Contra	0	266	63.6	1.60	10	162	33.1	1.28	6
	50	275	57.9	1.45	7	173	33.1	1.33	−1
LED only	–	295	67.5	1.19	–	171	41.7	1.19	–

the younger group. This is demonstrated by computing the multisensory response enhancement (MRE) values, i.e., the percentage of response speed-up in the bimodal condition relative to the visual-only condition (see columns 5 and 8 in Table 2).

This pattern of results is consistent with the studies of Laurienti et al. (2006) and Peiffer et al. (2007) in finding larger multisensory interaction effects in the elderly group. Interestingly, the Peiffer et al. study, where (manual) reaction times were in approximately the same range as our saccadic reaction times, did not find any significant age differences for the unimodal stimulus conditions, whereas we found saccadic eye movements to the visual-only signal much slower for the elderly group compared to the young adults (293 ms vs. 171 ms on average). What remains unclear is whether this slowing is due to an age-related decrement in visual processing or in a more centrally located processing stage where integration of visual and auditory input is taking place. In an attempt to gain further evidence on this issue, a parametric version of the TWIN model was fitted to our data. We present a sketch of the model first.

4. Time-window-of-integration (TWIN) model

4.1. Assumptions and predictions

The anatomical separation of the afferent pathways for the visual and auditory modality suggests at least two serial stages of saccadic reaction time: an early, afferent stage of peripheral processing (*first stage*) followed by a compound stage of converging subprocesses (*second stage*). In the first stage, a race among the peripheral neural excitations in the visual and auditory pathways triggered by a crossmodal stimulus complex takes place. The second stage comprises neural integration of the input and preparation of an oculomotor response. It is hypothesized that crossmodal interaction manifests itself in an increase or decrease of second stage processing time. Moreover, in the redundant target paradigm, the first stage duration is determined by the time of the winner of the race, whereas in the focused attention task – the only case considered here – the first stage duration is determined by the time it takes to process the target stimulus.

Thus, the model retains the classic notion of a race mechanism as an explanation for crossmodal interaction (cf. Colonius & Arndt, 2001; Miller, 1982; Mordkoff & Yantis, 1991; Raab, 1962) but restricts it to the very first stage of stimulus processing. The assumption of only two stages is certainly an oversimplification. Note, however, that the second stage is defined by default: it includes all subsequent, possibly overlapping, processes that are not part of the peripheral processes in the first stage (for a similar approach, see Van Opstal & Munoz, 2004).

The TWIN model makes further specific assumptions about the temporal configuration needed for multisensory integration to occur (see also Colonius & Diederich 2004; Diederich & Colonius 2007a):

(1) *Time-window-of-integration assumptions*. In the focused attention paradigm, crossmodal interaction occurs only if (i) a non-target stimulus wins the race in the first stage, opening a "time window" such that (ii) the termination of the target peripheral process falls in the window. The duration of the "time window" is a constant.

This means that the winning non-target will keep the saccadic system in a state of crossmodal reactivity such that the upcoming target stimulus, if it falls into the time window, will trigger crossmodal interaction. In the case of the target being the winner, no discernible effect on saccadic reaction time is predicted, analogous to the unimodal situation. In both cases, however, the first stage is terminated by the peripheral visual process.

The window of integration acts as a filter determining whether the afferent information delivered from different sensory organs is registered close enough in time for crossmodal interaction to take place. Passing this filter is necessary for crossmodal interaction to occur. It is not a sufficient condition because interaction also depends on the spatial configuration of the stimuli. Rather than assuming the existence of a joint spatiotemporal window of integration permitting interaction to occur only for both spatially and temporally neighboring stimuli, the TWIN model allows for interaction to occur even for rather distant stimuli of different modalities, as long as they fall exactly within the time window.

(2) *Assumption of spatiotemporal separability*. The amount of interaction in second-stage processing time is a function of the spatial configuration of the stimuli, but it does not depend on their (physical) presentation asynchrony (SOA).

Interaction, if it occurs at all, will either be inhibition or facilitation depending on both target and non-target positions. Typically, any facilitation decreases with the distance between the stimuli. More specific hypotheses about the effect of the spatial configuration on the amount of interaction have been studied in Diederich and Colonius (2007b).

Due to its two-stage processing assumption, it is easy to derive (see Appendix A for details) that the expected amount of crossmodal interaction (ECI) in the TWIN model, defined as expected SRT in unimodal (target) trials minus expected SRT in crossmodal trials, is the product of two components: the probability of interaction to occur ($P(I)$, say) and the amount of interaction [in ms] (Δ, say):

$$\mathrm{ECI} \equiv \mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] - \mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}] = P(I) \cdot \Delta. \qquad (1)$$

Since, by the spatiotemporal separability assumption, the amount of interaction, Δ, does not depend on SOA, the dependence of the expected amount of crossmodal interaction, ECI, on SOA is solely determined by how the probability of interaction, $P(I)$, depends on SOA. Consider two extreme SOA conditions: when the non-target is presented very late relative to the target (large positive SOA), its chances of winning the race against the target and thus opening the window of integration are small. On the other hand, if it is presented very early (large negative SOA), it is likely to win the race and to open the window, but the window may close before the arrival of the target. Again, the probability of interaction, $P(I)$, is small. Therefore, the largest effects are expected for some mid-range SOA values.

The effect of increasing the window size is also easy to appreciate intuitively: whenever the window has been opened, the chances that the target will arrive and interaction does occur will be the higher the longer the window stays open. That is, the expected amount of crossmodal interaction will be a monotonically increasing function of window size.

4.2. *Fitting the model to the data*

In order to provide a quantitative fit we need to specify further the probability distributions for the processing times in the first stage. For simplicity, we assume an exponential distribution for the peripheral processing times V for a visual target and A for an auditory non-target, respectively, with parameters λ_V and λ_A. With exponential distributions, the expected response time for the crossmodal and unimodal conditions becomes

$$\mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}] = \frac{1}{\lambda_V} + \mu - P(I) \cdot \Delta \qquad (2)$$

$$\mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] = \frac{1}{\lambda_V} + \mu \qquad (3)$$

respectively. Here, the mean of second stage processing time (without interaction occurring) is taken to be μ, where we need not specify the distribution as long as predictions are restricted to the expected values of SRT.[1] Since the amount of interactions is allowed to depend on the spatial configuration, we introduce two parameters, Δ_{ipsi} and Δ_{contra}. The final parameter to be estimated is size of the time window, ω.

Table 3

Estimated parameters and best-fit χ^2 for both groups under the following parameter restrictions: $\lambda_V, \lambda_A, \mu > 0$

Age group	$1/\lambda_V$	$1/\lambda_A$	μ	Δ_i	Δ_c	ω	χ^2_{est}	χ^2_{pred}
65–75	84	98	200	104	38	450	2.10	4.96
20–22	48	18	131	44	25	275	0.001	3.29

Thus, for each age group six parameters ($\lambda_V, \lambda_A, \mu, \Delta_{\mathrm{ipsi}}, \Delta_{\mathrm{contra}}, \omega$) were estimated from their eight observed mean SRT values for visual–auditory stimulus pairs presented ipsi- and contralaterally in 4 SOA steps. To check for the robustness of these estimates, only half of the data (the odd trial numbers) were used for the estimation procedure and the fit was assessed in predicting the means computed on the other half (the even trial numbers). Importantly, mean SRT to the unimodal stimulus (LED) was not used in the estimation procedure but, instead, predicted from the estimated model providing an additional check of the model.[2]

The model fit to the mean SRT is presented in Fig. 2, panel (A). The predicted probabilities of interaction ($P(I)$) are found in panel (B). Note that the probabilities cannot be observed directly: multiplying these estimates by the estimated Δ values results in the (observable) difference between uni- and bimodal average saccadic reaction time, according to Eq. (1).

From visual inspection, TWIN gives an excellent account of the data for both groups. Standard errors are too small to show up in the graphs. The estimated parameters for both groups are found in Table 3. Notably, the two age groups differ systematically in all parameters. Mean peripheral processing time for visual and auditory stimuli, $1/\lambda_V$ and $1/\lambda_A$, respectively, and mean central processing time, μ, are longer for the older participants than for the younger ones. The amount of interaction for ipsi- and contralaterally presented bimodal stimuli (Δ_{ipsi} and Δ_{contra}) is larger for the elderly. Finally, the window-of-integration width, ω, is also larger for the elderly.

Despite these differences in the parameters, the predictions reflect the overall similarity of the response time patterns between the age groups: the increase of mean SRT as a function of SOA, the faster mean SRT for ipsi- than for contralaterally presented bimodal stimulus combinations, and the overall bimodal enhancement. Note, however, that the younger participants are not only faster than the elderly under all conditions, but the increase of SRT over SOAs is also less steep. As suggested by a reviewer, this may in part be a "floor" effect in that their reaction times cannot drop much below the 140 ms found for the fastest conditions. Within the TWIN model explanation, the probability of interaction decreases much faster for the younger participants while the amount of interaction (Δ) remains constant over SOA: the probability is very high (close to 1) for large negative SOAs (-100 and -50) and then decreases sharply falling to almost zero for positive SOA of 50. For the elderly participants, the probability of interaction is lower for the negative SOA (around 0.7 for SOA $= -100$) and decreases linearly with SOA. For SOA $= 50$ the probability of an interaction is still at around 0.3.

---

[1] If the entire distribution of SRT is considered, a possible choice for the second stage is the normal distribution resulting in a mixture of ex-Gaussian distributions for the bimodal conditions (see Colonius & Arndt, 2001).

[2] It should be noted here that more elaborate ways of model testing are possible. However, given the low number of data points relative to the number of parameters, this study was not meant to be a stringent test of the TWIN model assumptions. Such tests have been performed in previous investigations with large numbers of stimulus onset asynchrony values and various spatial configurations (Diederich & Colonius, 2007a,b, 2008).

Note that there exists also the possibility that the auditory non-target, when presented early enough, may act as a spatially unspecific cue for the upcoming visual target. In a recent extension of TWIN (Diederich & Colonius, 2008), we have shown that this effect can be separated quantitatively from the multisensory integration effect proper. Given the restricted SOA range in the current study, we refrained from considering this extended model version.

5. Summary and discussion

Crossmodal visual–auditory interaction effects could be observed in a focused attention paradigm for two different age groups, young adults (aged 20–22 years) and older adults (aged 65–75 years). The latter were slower under all experimental conditions but showed greater multisensory enhancement than younger adults. These results are in line with results by Laurienti et al. (2006) and Peiffer et al. (2007) using manual reaction times in a redundant target paradigm, except that the latter did not find any significant age differences for the unimodal stimulus conditions.

The aim of the current study was to quantify the contributions of the sensory and cognitive processes possibly responsible for the difference in the younger and older adults by fitting the observed data to the TWIN model. TWIN assumes two serial interdependent processing stages: a first stage encompassing all stimulus-triggered parallel and independent peripheral processes, and a second stage comprising central processes including neural integration and oculomotor response initiation. It assumes the existence of a time window into which termination of the peripheral processes need to fall for crossmodal integration to occur.

After specifying the probability distributions for the peripheral processes, the relevant parameters were estimated from the observed SRT data separately for both age groups. These parameter estimates point to the following interpretation. First, peripheral processing takes substantially longer for the older participants than for the younger ones (84 ms for visual and 98 ms for auditory stimuli compared to 48 ms and 18 ms for the younger group, respectively).

Second, estimating the width of the time-window-of-integration yielded 450 ms for the older participants and 275 ms for the younger group. This is in accordance with Laurienti et al. (2006) who concluded from their race model analysis that older adults have a broader temporal window of integration than younger adults. Our analysis here reveals, in addition, that for the elderly the probability of integration ($P(I)$) is smaller (for negative SOAs) even with a broader window. However, note that this probability is not only determined by the window width but also by the processing times for the peripheral processes. That is, as peripheral processing seems to slow down with age, the probability of integrating information from different sources declines because the time to process visual and auditory sensory information is more variable and less likely to terminate within the time window. Thus, interestingly, in TWIN this reduced integration capability in the elderly is a direct consequence of the peripheral slowing and not an indication of a general cognitive slowing. A broader temporal window, it seems, can only partially compensate for this. From a neurophysiological point of view, the effect of the auditory non-target winning the peripheral race may correspond to an early inhibition of the activity of the fixation neurons in superior colliculus and/or of the omnipause neurons in the brainstem (e.g., Wurtz, Basso, Paré, & Sommer, 2000). In the elderly then the likelihood for this early inhibition seems to decline mainly because of a lengthening of the peripheral processes.

Finally, the predicted amount of neural enhancement for ipsi- and contralaterally presented stimuli (Δ_{ipsi} and Δ_{contra}, respectively) was found larger for the elderly (104 ms and 38 ms, respectively) compared to the corresponding values for the younger group (44 ms and 25 ms, respectively).

Note that treating the decline in peripheral processing as equivalent to a reduction in stimulus intensity here raises an analogy with the principle of “inverse effectiveness” (e.g., Stein & Meredith, 1993) according to which multisensory enhancement for weaker stimuli tends to be larger than for more intense stimuli. On the other hand, it is not obvious whether increasing stimulus intensity would lead to a pattern of results in the elderly similar to what has been observed in the younger group. This will be the case only if a reduction of mean SRT is accompanied by a corresponding reduction in SRT variability. Since the TWIN model (with its exponential distribution assumption) predicts just this, an experiment with varying stimulus intensity levels for the age groups would constitute an interesting goal for future investigations.

In conclusion, the following picture emerges: (i) there is an undisputable slowing of the peripheral sensory processing in the elderly, (ii) as a result of this slowing, the probability of integration is smaller in the elderly even with a wider time window of integration, and (iii) multisensory integration, if it occurs, manifests itself in larger neural enhancement in the elderly; however, because of (ii), on average the integration effect is not large enough to compensate for the peripheral slowing in the elderly.

Acknowledgment

This research was supported by grants from Deutsche Forschungsgemeinschaft Di 506/8-1 and /-3 and Sonderforschungsbereich/Transregio 31 “Active Listening”, Project B4.

Appendix A

A.1. Formal specification of TWIN

The race in the first stage of the model is made explicit by assigning independent nonnegative random variables V and A to the peripheral processing times for the visual target and auditory non-target stimulus, respectively. With τ as SOA value and ω_{I} as integration window width parameter, the time window of integration assumption is equivalent to the (stochastic) event I, say

$$I = \{A + \tau < V < A + \tau + \omega\}.$$

Thus, the probability of integration to occur, $P(I)$, is a function of both τ and ω, and it can be determined numerically once the distribution functions of A and V have been specified (see below).

The next step is to compute expected reaction time for the unimodal and crossmodal conditions. From the two-stage assumption, total reaction time in the crossmodal condition can be written as a sum of two random variables:

$$\text{RT}_{\text{crossmodal}} = S_1 + S_2, \tag{4}$$

where S_1 and S_2 refer to the first and second stage processing time, respectively. For the expected saccadic reaction time in the crossmodal condition then follows:

$$\begin{aligned}
\text{E}[\text{RT}_{\text{crossmodal}}] &= \text{E}[S_1] + \text{E}[S_2] \\
&= \text{E}[S_1] + P(I)\text{E}[S_2|I] + (1 - P(I))\text{E}[S_2|\text{not-}I] \\
&= \text{E}[S_1] + \text{E}[S_2|\text{not-}I] - P(I)(\text{E}[S_2|\text{not-}I] - \text{E}[S_2|I]),
\end{aligned}$$

where $\text{E}[S_2|I]$ and $\text{E}[S_2|\text{not-}I]$ denote the expected second stage processing time conditioned on interaction occurring (I) or not occurring (not-I), respectively. Setting

$$\Delta \equiv \mathrm{E}[S_2|\text{not-}I] - \mathrm{E}[S_2|I]$$

this becomes

$$\mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}] = \mathrm{E}[S_1] + \mathrm{E}[S_2|\text{not-}I] - P(I) \cdot \Delta. \tag{5}$$

In the unimodal condition, no integration is possible. Thus,

$$\mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] = \mathrm{E}[S_1] + \mathrm{E}[S_2|\text{not-}I],$$

and we arrive at the simple product rule for crossmodal interaction (CI)

$$\mathrm{ECI} \equiv \mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] - \mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}] = P(I) \cdot \Delta. \tag{6}$$

Δ takes on positive or negative values (or zero) depending on whether multisensory integration has a facilitative or inhibitory effect.

Exponential distribution are postulated, for simplicity, for the peripheral processing time V for a visual target and A for an auditory non-target with parameters λ_{V} and λ_{A}, respectively. That is:

$$f_{\mathrm{V}}(t) = \lambda_{\mathrm{V}}\, \mathrm{e}^{-\lambda_{\mathrm{V}} t},$$
$$f_{\mathrm{A}}(t) = \lambda_{\mathrm{A}}\, \mathrm{e}^{-\lambda_{\mathrm{A}} t},$$

for $t \geq 0$, and $f_{\mathrm{V}}(t) = f_{\mathrm{A}}(t) \equiv 0$ for $t < 0$. The corresponding distribution functions are referred to by $F_{\mathrm{V}}(t)$ and $F_{\mathrm{A}}(t)$.

In order to compute

$$\begin{aligned} P(I) &= P(A + \tau < V < A + \tau + \omega) \\ &= \int_0^{\infty} f_{\mathrm{A}}(a)\{F_{\mathrm{V}}(a + \tau + \omega) - F_{\mathrm{V}}(a + \tau)\}\, \mathrm{d}a, \end{aligned}$$

it is necessary to distinguish three cases for the magnitude of $\tau + \omega$ resulting in the following expressions:

(i) $\tau < \tau + \omega < 0$

$$P(I) = \frac{\lambda_{\mathrm{V}}}{\lambda_{\mathrm{V}} + \lambda_{\mathrm{A}}}\, \mathrm{e}^{\lambda_{\mathrm{A}} \tau}(-1 + \mathrm{e}^{\lambda_{\mathrm{A}} \omega});$$

(ii) $\tau < 0 < \tau + \omega$

$$P(I) = \frac{1}{\lambda_{\mathrm{V}} + \lambda_{\mathrm{A}}}\, \{\lambda_{\mathrm{A}}(1 - \mathrm{e}^{-\lambda_{\mathrm{V}}(\omega+\tau)}) + \lambda_{\mathrm{V}}(1 - \mathrm{e}^{\lambda_{\mathrm{A}} \tau})\};$$

(iii) $0 < \tau < \tau + \omega$

$$P(I) = \frac{\lambda_{\mathrm{A}}}{\lambda_{\mathrm{V}} + \lambda_{\mathrm{A}}}\, \{\mathrm{e}^{-\lambda_{\mathrm{V}} \tau} - \mathrm{e}^{-\lambda_{\mathrm{V}}(\omega+\tau)}\}.$$

A.2. Model fit procedure

Parameters were estimated by minimizing the Pearson χ^2 statistic

$$\chi^2 = \sum_{n=1}^{4} \sum_{j=1}^{2} \left(\frac{\overline{\mathrm{SRT}}(j, n) - \widehat{\mathrm{SRT}}(j, n)}{\sigma_{\overline{\mathrm{SRT}}(j,n)}} \right)^2 \tag{7}$$

using the FMINSEARCH routine of MATLAB. Here $\overline{\mathrm{SRT}}(j, n)$ and $\widehat{\mathrm{SRT}}(j, n)$ are, respectively, the observed and the fitted values of the mean SRT to visual–auditory stimuli presented in spatial positions (ipsilateral, $j = 1$; contralateral, $j = 2$) with SOA (referred to by n to 4); $\sigma_{\overline{\mathrm{SRT}}(j,n)}$ are the respective standard errors.

For the estimation routine λ_{V}, λ_{A}, μ, and ω were restricted to positive real numbers.

References

Calvert, G., Spence, C., & Stein, B. E. (2004). *Handbook of multisensory processes*. Cambridge, MA: MIT Press.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*, 67–83.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, *34*, 253–275.

Colonius, H., & Arndt, P. (2001). A two-stage model for visual–auditory interaction in saccadic latencies. *Perception & Psychophysics*, *63*, 126–147.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience*, *16*(6), 1000–1009.

Colonius, H., & Diederich, A. (2006). Race model inequality: Interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*(1), 148–154.

Diederich, A., & Colonius, H. (2004). Modeling the time course of multisensory interaction in manual and saccadic responses. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes*. Cambridge, MA: MIT Press.

Diederich, A., & Colonius, H. (2007a). Why two "distractors" are better then one: Modeling the effect on non-target auditory and tactile stimuli on visual saccadic reaction time. *Experimental Brain Research*, *179*, 43–54.

Diederich, A., & Colonius, H. (2007b). Modeling spatial effects in visual-tactile saccadic reaction time. *Perception & Psychophysics*, *69*(1), 56–67.

Diederich, A., & Colonius, H. (2008). Crossmodal interaction in saccadic reaction time: Separating multisensory from warning effects in the time window of integration model. *Experimental Brain Research*, *186*, 1–22.

Helfer, K. S. (1998). Auditory and auditory–visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology*, *9*, 234–242.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. London: Wiley.

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, *27*, 1155–1163.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Miller, J. O. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, *14*, 115–128.

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 520–538.

Peiffer, A. M., Mozolica, J. L., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *NeuroReport*, *18*(10), 1077–1081.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Science*, *24*, 574–590.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.

Todd, J. W. (1912). Reaction to multiple stimuli. In R. S. Woodworth (Ed.), *Archives of psychology, No. 25. Columbia contributions to philosophy and psychology*. New York: The Science Press.

Van Opstal, A. J., & Munoz, D. P. (2004). Auditory–visual interactions subserving primate gaze orienting. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes*. Cambridge, MA: MIT Press.

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In: K. R. Boff, L. Kaufmann, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. 1. Sensory processes and perception* (pp. 25-1–25-36). New York: Wiley.

Woods, T. M., & Recanzone, G. H. (2004). Visually induced plasticity of auditory spatial perception in macaques. *Current Biology*, *14*, 1559–1564.

Wurtz, R. H., Basso, M. A., Paré, M., & Sommer, M. A. (2000). The superior colliculus and the cognitive control of movement. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., Vol. XXI, No. 8, pp. 573–588). Cambridge, MA: MIT Press.